Processes maintaining tropopause sharpness in numerical models

L. Saffin [1], S. L. Gray [1], J. Methven [1] and K. D. Williams [2]

[1]Department of Meteorology, University of Reading, Reading, UK
[2]Met Office, Exeter, UK

Key Points:

- Tracers of potential vorticity are used to investigate the evolution of tropopause sharpness and reasons for systematic model error
- The tropopause smooths due to an imbalance between parametrized processes sharpening and the advection scheme smoothing the tropopause
- Sharpening is weaker and advective smoothing is more rapid in ridges compared to troughs

Corresponding author: Leo Saffin, `L.Saffin@pgr.reading.ac.uk`

Abstract

Recent work has shown that the sharpness of the extratropical tropopause declines with lead time in numerical weather prediction models, indicating an imbalance between processes acting to sharpen and smooth the tropopause. In this study the systematic effects of processes contributing to the tropopause sharpness are investigated using daily initialised forecasts run with the Met Office Unified Model over a three-month winter period. Artificial tracers, each forced by the potential vorticity tendency due to a different model process, are used to separate the effects of such processes. The advection scheme is shown to result in an exponential decay of tropopause sharpness towards a finite value at short lead times with a timescale of 20-24 hours. The systematic effect of non-conservative processes is to sharpen the tropopause, consistent with previous case studies. The decay of tropopause sharpness due to the advection scheme is stronger than the sharpening effect of non-conservative processes leading to a systematic decline in tropopause sharpness with forecast lead time. The systematic forecast errors in tropopause-level potential vorticity are comparable to the integrated tendencies of the parametrized physical processes suggesting that the systematic error in tropopause sharpness could be significantly reduced through realistic adjustments to the model parametrization schemes.

1 Introduction

A distinct feature of the extratropical atmosphere is the sharp contrast between the troposphere and the stratosphere: the tropopause. The thermal tropopause is defined as the height at which the vertical lapse rate transitions from tropospheric values to stratospheric values. Composites of radiosonde data in height relative to the thermal tropopause show a shallow static stability maximum above the tropopause known as the tropopause inversion layer (TIL) [*Birner et al.*, 2002] emphasising that the vertical transition in lapse rate is sharp. The dynamical tropopause defines the boundary between the troposphere and stratosphere as a value of Ertel potential vorticity (PV) between the tropospheric values and stratospheric values. Since PV is conserved for adiabatic and frictionless motion [*Ertel*, 1942], the dynamical tropopause emphasises that the tropopause behaves almost like a material surface with exchange of mass between the stratosphere and troposphere only enabled by diabatic processes (including small-scale mixing).

Since both potential temperature (θ) and PV are conserved for adiabatic and frictionless motion, the large-scale dynamics of the midlatitude atmosphere are compactly described by maps of PV on isentropic (constant θ) surfaces [*Hoskins et al.*, 1985] where the tropopause is seen as a narrow region of strong isentropic gradients of PV separating the high PV stratospheric air and the low PV tropospheric air. The strong isentropic PV gradient at the tropopause, coinciding with the midlatitude jet, acts as a waveguide for Rossby-waves [*Hoskins and Ambrizzi*, 1993; *Martius et al.*, 2010]. Rossby waves can be an important source of predictability in medium-range forecasting [*Grazzini and Vitart*, 2015] and are crucial to accurately representing longer time-scale processes [*Palmer et al.*, 2008].

The isentropic tropopause PV gradient decreases systematically with forecast lead time in current numerical weather prediction (NWP) models [*Gray et al.*, 2014]. Rossby wave propagation depends on to the isentropic PV gradient: a weaker tropopause PV gradient both reduces jet speed and weakens the upstream propagation rate of Rossby-waves. *Harvey et al.* [2016] showed that the two effects cancel at first order but at second order the reduction in jet speed is greater, giving a net reduction in phase speed. They estimated that the smoother isentropic PV gradients seen in NWP forecasts compared to analyses would produce a phase error in Rossby waves of 400 km over 5 days.

The reduction of the tropopause PV gradient with forecast lead time indicates that there is a net imbalance in the processes modifying the tropopause PV gradient. The purpose of this study is to quantify the systematic effects of different processes within an NWP model contributing to the tropopause PV gradient and so provide a method for model developers to link systematic forecast errors with the physical processes responsible. The systematic difference between forecasts and analyses is equivalent to the systematic imbalance between model processes at short lead times, and attributing tendencies to individual model processes can give insight into the origin of model imbalances [*Klinker and Sardeshmukh*, 1992; *Rodwell and Palmer*, 2007]. In this study we are interested in the "initial tendencies" contributing to the tropopause PV gradient; however, a limitation of the initial tendencies method is the potential for advection to dominate the tendencies due to the Eulerian frame. Hence, artificial tracers, described as PV tracers, are used to accumulate tendencies of PV from individual model processes in a Lagrangian frame. PV tracers allow us to better quantify the integrated effect of different processes on PV following air masses, following the method of *Davis et al.* [1993] and *Saffin et al.* [2016].

The structure of this paper is as follows. A brief review of the key processes affecting the tropopause sharpness is given in section 2. The setup of the forecasts analysed including the online integration of PV tracers is given in section 3 as well as an objective definition of ridges and troughs used in compositing the forecasts. Section 4 describes the results. The key conclusions and discussion of results are presented in section 5.

2 Processes affecting tropopause sharpness

From previous studies, three key processes affecting tropopause sharpness have been identified: vortex stripping, radiative cooling and latent heating enhanced ascent (i.e. warm conveyor-belts (WCBs)) have significant effects on the midlatitude tropopause. In this study the relative contributions of these processes are quantified using daily forecasts over a winter season.

Vortex stripping describes a process in which sharp gradients in vorticity are generated from an initially smooth vorticity distribution in two-dimensional fluids [*Legras and Dritschel*, 1993]. Using an isentropic single-layer quasi-geostrophic model, *Ambaum* [1997] showed that the two-dimensional vortex stripping motion of baroclinic eddies is the essential process for forming and maintaining a sharp tropopause PV gradient. Results of three-dimensional simulations have shown that layerwise horizontal vortex stripping in isentropic layers can also result in sharp vertical PV gradients [*Haynes et al.*, 2001] and a TIL [*Son and Polvani*, 2007; *Wang and Geller*, 2016]. The general action of vortex stripping can be described as air being stirred on either side of the tropopause without stirring across the tropopause which acts as a transport barrier. We can approximately consider that the stirring results in a three-component fluid on an isentrope with high-PV stratospheric air around the poles and low-PV tropospheric air equatorward, separated by a region of intermediate PV: the tropopause. The regions of intermediate PV are drawn away from the tropopause by the eddies on either side of the tropopause. The intermediate PV is then stretched out into filaments. As the filaments stretch out they are broken up by small-scale mixing and gradually dissipated. The result is that the PV gradient at the tropopause has been enhanced by removing the intermediate PV air and bringing high and low PV air closer together. At longer time scales small-scale mixing will eventually dominate resulting in a uniform PV distribution; a key process for maintaining the tropopause sharpness in idealized simulations is the inclusion of a thermal relaxation towards a state with a smooth equator-to-pole PV gradient, as an idealized representation of other diabatic processes, which acts to maintain the contrast between the high-PV stratospheric air and the low-PV tropospheric air. The result is a dynamical equilibrium between thermal relaxation and vortex stripping [*Ambaum*, 1997; *Haynes et al.*, 2001].

The effects of diabatic processes on the tropopause are more complicated than thermal relaxation: *Forster and Wirth* [2000] showed that radiative cooling could directly enhance the PV contrast across filaments of PV provided the vorticity was sufficiently large and *Randel et al.* [2007] showed that radiative cooling provides a significant contribution to the strength of the TIL. The dominant contribution to the direct effect of radiation on the tropopause is long-wave cooling from water vapour [*Forster and Wirth*, 2000; *Ferreira et al.*, 2016]: the moister troposphere cools more rapidly than the drier stratosphere with the most efficient cooling just below the dry layer resulting in a gradient of diabatic heating and positive PV tendencies across the humidity gradient. The presence of clouds will modify the profile of radiation and, as a result, the PV tendencies. The addition of clouds below the tropopause acts to focus the maxima in radiative cooling at the cloud top [*Cau et al.*, 2005], resulting in a sharper gradient in diabatic heating rate and a stronger and more localised dipole of PV tendencies, positive above the cloud and negative below.

Latent heating in WCBs has been shown to affect the tropopause. WCBs are air streams associated with extratropical cyclones which transport air upwards and polewards [*Harrold*, 1973]. A WCB airstream can be identified as a coherent ensemble of trajectories ascending 600 hPa in 48 hours following *Wernli and Davies* [1997]. WCBs transport moist low-PV air from the boundary layer to the upper troposphere [*Wernli and Davies*, 1997] and the outflow can have large impacts on the tropopause and subsequent Rossby wave propagation [*Riemer and Jones*, 2010; *Grams et al.*, 2011]. Latent heating has a large effect on WCB evolution: air parcels typically experience a net heating of ≈20 K [*Madonna et al.*, 2014b] mainly associated with condensation at low-levels and depositional growth of snow at upper-levels [*Joos and Wernli*, 2012]. *Schemm et al.* [2013] showed that a dry simulation produced a weaker WCB and as a result slower development of a downstream cyclone when compared with a moist simulation. In terms of PV, air parcels experience positive PV tendencies below the maximum in latent heating rates and negative PV tendencies above. WCB climatologies have found the net change in PV between the inflow and outflow of WCB trajectories to be close to zero [*Madonna et al.*, 2014b]. *Methven* [2015] used a Kelvin's circulation argument to outline the conditions under which the PV of the inflow is expected to match that of the outflow.

Chagnon et al. [2013] showed that the combined effect of long-wave radiation and WCBs gave a dipole of diabatically-generated PV that enhanced the tropopause PV gradient. *Chagnon et al.* [2013] also argued that the transport of moisture by the WCB would enhance the effects of long-wave radiation. *Kunkel et al.* [2016] showed similar results for the TIL: long-wave radiation strengthened the TIL and transport of moisture to the tropopause results in a more rapid formation of the TIL. However, these results are limited to case studies [*Chagnon et al.*, 2013; *Chagnon and Gray*, 2015] and idealised simulations [*Kunkel et al.*, 2016]. This study instead quantifies the systematic effects of physical processes on the tropopause over a season of forecasts with an NWP model.

3 Methods

The data analysed in this paper are from a set of forecasts run with the NAE (North Atlantic and European) configuration of the MetUM version 7.3 (section 3.1). The online integration of PV tracers with the MetUM is described in section 3.2. The data output from the forecasts has been composited separately for ridges and troughs; a new diagnostic for ridges and troughs is described in section 3.3.

3.1 Forecasts with the MetUM

The Met Office Unified Model (MetUM) is an operational NWP model. The dynamical core of the MetUM version used here approximates a two time level, semi-implicit, semi-Lagrangian solution to the nonhydrostatic, deep atmosphere equations [*Davies et al.*, 2005]. The variables in the MetUM are placed on a C-grid [*Arakawa and Lamb*, 1977] with Charney-Phillips staggering in the vertical [*Charney and Phillips*, 1953] using a terrain-following, height-based, coordinate that gradually flattens at higher altitudes [*Davies et al.*, 2005]. The MetUM contains various parametrizations to account for physical processes that are either not resolved or not represented within the dynamical core: radiation [*Edwards and Slingo*, 1995]; microphysics [*Wilson and Ballard*, 1999]; orographic [*Webster et al.*, 2003] and non-orographic [*Scaife et al.*, 2002] gravity-wave drag; convection [*Gregory and Rowntree*, 1990]; and turbulent mixing [*Lock et al.*, 2000].

A forecast was initialised for each day in the three-month winter period from 1 November 2013 to 31 January 2014 (a total of 92 forecasts). The forecasts were run using the limited area NAE configuration (see Fig. 1 for domain extent). The period was chosen to use the most recent available analyses for the NAE configuration. The Met Office phased out operational use of the NAE domain beyond January 2014 which is why we have used November instead of February for our "winter" season. The NAE domain has 0.11° horizontal grid spacing and uses a rotated pole to center the domain on the equator giving an, approximately uniform, 12-km grid spacing. We use 70 non-uniformly spaced vertical model levels up to 80 km with a 5-minute timestep. The initial conditions used are from operational NAE analyses, and boundary conditions are given by operational runs of the global model for the same start time using the method described in *Davies* [2014]. Each forecast was initialised at 00 UTC and run for 2.5 days to give an overlap between forecasts.

3.2 PV Tracers

A set of PV tracers were integrated online with each of the forecasts to quantify the effects of the various processes in the MetUM. The method is based on *Davis et al.* [1993] and was first applied to the MetUM by *Gray* [2006]. The method works by partitioning and integrating PV. Following an air parcel, PV is modified by diabatic and frictional processes,

$$\frac{Dq}{Dt} = \frac{1}{\rho}\boldsymbol{\zeta} \cdot \nabla\dot{\theta} + \frac{1}{\rho}\nabla\dot{\theta} \cdot \nabla \times \mathbf{F}, \tag{1}$$

where $q = \frac{1}{\rho}\boldsymbol{\zeta} \cdot \nabla\theta$ is Ertel PV [*Ertel*, 1942], ρ is density, $\boldsymbol{\zeta}$ is the absolute vorticity vector, $\dot{\theta}$ is the diabatic heating tendency and $\mathbf{F}$ is friction. The general method is to integrate the tendency of PV along trajectories over a forecast interval T:

$$\int_{t_0}^{t_0+T} \frac{Dq}{Dt} dt = \int_{t_0}^{t_0+T} S dt, \tag{2}$$

where t_0 is the forecast start time and S represents the right hand side of Eq. 1, which can be partitioned by process ($S = \sum S_i$) resulting in a set of PV tracers ($\sum q_i$) where

$$q_i = \int_{t_0}^{t_0+T} S_i dt. \tag{3}$$

In the MetUM the PV tendencies are derived from the parametrization schemes resulting in a set of physics tracers ($\sum q_{phys}$): short-wave radiation (q_{sw}), long-wave radiation (q_{lw}), microphysics (q_{mic}), gravity-wave drag (q_{gwd}), convection (q_{con}) and turbulent mixing (q_{tm}). There is also a term for "cloud re-balancing"; however, this is not shown in later composites as it is negligible. For the initial conditions, an advection-only PV tracer (q_{adv}) is set equal to the full PV with all other tracers set to zero. This initiation is also applied in the lateral boundaries of the limited area domain at every timestep. Every timestep each PV tracer is incremented by its respective PV tendency (zero for q_{adv}) and advected by the semi-Lagrangian advection scheme of the model.

The "dynamics-tracer inconsistency" diagnostic (ε_I) of *Saffin et al.* [2016] is also included. The dynamics-tracer inconsistency quantifies the difference between the PV tendency diagnosed by the dynamical core and the PV tendency diagnosed by tracer advection of PV. In a single timestep there will be a change in PV due to the dynamical core modifying prognostic variables,

$$\Delta q_{dyn} = q(\mathbf{X}^{(1)}) - q(\mathbf{X}^{(0)}), \tag{4}$$

where $q(\cdot)$ represents a calculation of PV from the prognostic variables in the model at the start of the timestep ($\mathbf{X}^{(0)}$) and after the dynamics terms are calculated ($\mathbf{X}^{(1)}$). We can also calculate a change in PV due to tracer advection,

$$\Delta q_{tracer} = q(\mathbf{X}^{(0)})_d - q(\mathbf{X}^{(0)}), \tag{5}$$

where the subscript '*d*' denotes evalution at departure points in the MetUM's semi-Lagrangian method (i.e. tracer advection). The difference between these two changes gives

$$\Delta\varepsilon_I = \Delta q_{dyn} - \Delta q_{tracer} = q(\mathbf{X}^{(1)}) - q(\mathbf{X}^{(0)})_d. \tag{6}$$

This difference is calculated at every timestep and accumulated as an additional tracer (ε_I) in the same way as the physics tracers (i.e. Eq. 3). *Saffin et al.* [2016] showed that ε_I is an important component of a model PV budget which can be desirable to minimize (e.g. *Whitehead et al.* [2015]); however, exact conservation of PV is not necessarily a desirable property of a dynamical core because the cascade to smaller scales will be blocked at the grid-scale [*Thuburn*, 2008].

The final result is a budget for the Lagrangian change in PV,

$$q - q_{adv} = \sum q_{phys} + \varepsilon_I + \varepsilon_r, \tag{7}$$

where ε_r is calculated as a residual. The residual is a result of advecting multiple PV tracers with an imperfect advection scheme as well as any missing terms when summing increments over a timestep. The residual was shown to be more than an order of magnitude smaller than the dominant physics PV tracers by *Saffin et al.* [2016].

3.3 Objective definition of ridges and troughs

The results in this study are tropopause-relative composites produced over ridges and troughs separately. The expectation is that there will be significant differences in the behaviour of physical processes in ridges and troughs. For example, we might expect stronger effects of radiation in troughs due to a lower tropopause meaning more moist and cloudy air below the tropopause (e.g. *Cavallo and Hakim* [2009]), whereas we might associate ridges more with the strongly ascending WCB outflows. There are also differences in the structure of ridges and troughs purely due to the balanced dynamics [*Wirth*, 2001]. In this section a new diagnostic approach for dividing regions into ridges and troughs is described.

The diagnostic extends *Gray et al.* [2014] where the position of the tropopause is compared with an "equivalent latitude" (to be defined below). *Gray et al.* [2014] identify the location of the tropopause with a single contour of PV on 320 K: anywhere the contour is poleward of its equivalent latitude is a ridge and anywhere the contour is equatorward is a trough. *Hoskins and Berrisford* [1988] introduced maps of θ on the tropopause as a useful overview of multiple isentropic PV maps, where a value of 2 PVU (1 PVU = 10^{-6} m^2 s^{-1} K kg^{-1} [*Hoskins et al.*, 1985]) is typically used to define the tropopause. An isopleth of PV on a θ surface is the same as an isopleth of θ on a PV surface; therefore, a map of θ on the 2-PVU surface is equivalent to identifying the 2-PVU tropopause on every isentrope that intersects it. An exception is that the 2-PVU surface can fold so that the 2-PVU surface can be crossed multiple times on a vertical profile above some geographical locations. At any geographical location where the PV surface is folded we choose the highest value of θ. Ridges and troughs are then defined as anomalies of θ on the 2-PVU surface relative to a zonally symmetric background state:

$$\theta' = \theta(\lambda, \phi, q{=}2) - \theta_b(\phi, q{=}2), \tag{8}$$

where $\theta(\lambda, \phi, q{=}2)$ is the forecast θ as a function of longitude (λ) and latitude (ϕ) on the 2-PVU surface (q=2) and $\theta_b(\phi, q{=}2)$ is a zonally symmetric background state. A gridpoint is defined as being in a ridge or trough by a positive or negative value of θ' respectively.

The background state used here is defined by adiabatic rearrangement of PV to a zonally symmetric state [*Methven and Berrisford*, 2015]: for each PV contour on each isentropic surface an equivalent latitude (ϕ_e) is defined as the latitude circle that encloses the same mass and circulation as the PV contour in the full (3D) state. The method of *Methven and Berrisford* [2015] calculates a set of equivalent latitudes as a function of PV value on isentropic surfaces $\phi_e(\theta, q)$ at six-hourly intervals from ERA-Interim data [*Dee et al.*, 2011]. Figure 1a shows the range of $\phi_e(\theta, q{=}2)$ for the three-month forecast period with the first timestep of each month overplotted to highlight the instantaneous structure.

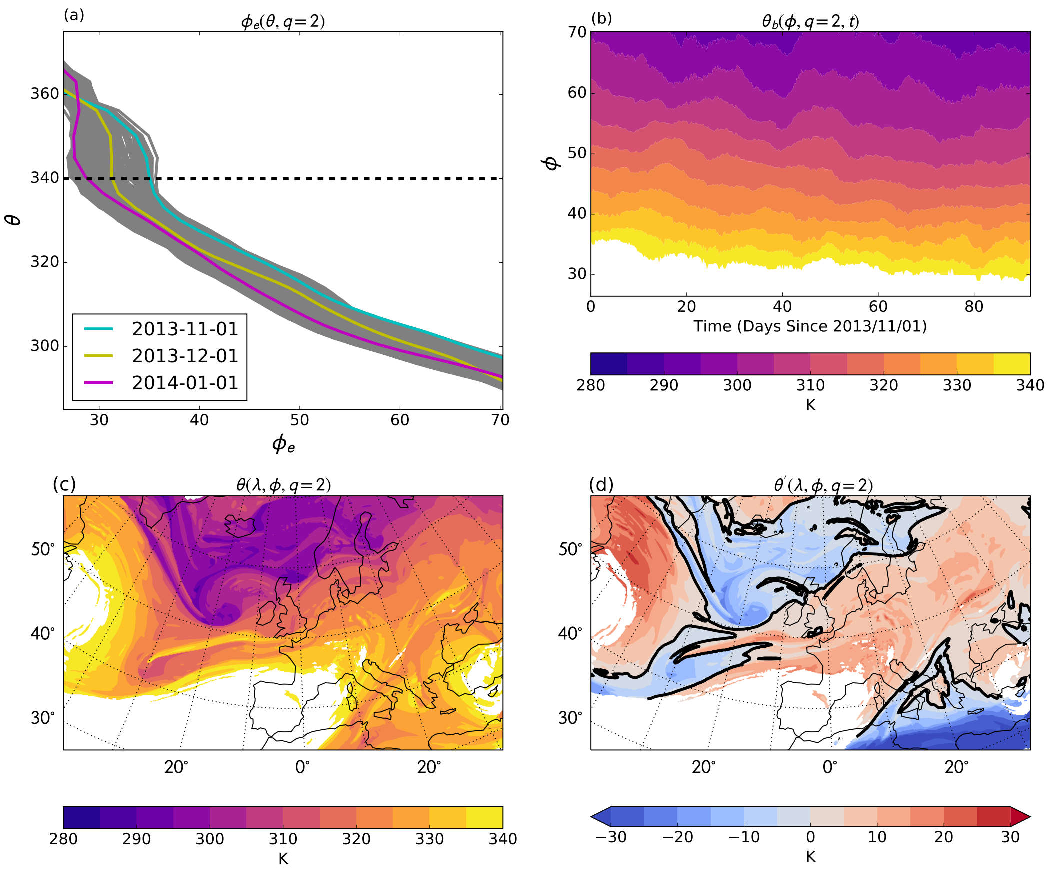

Figure 1. The background state on the 2-PVU surface used to diagnose ridges and troughs for the three-month forecast period. (a) $\phi_e(\theta, q{=}2)$, gray shows the range of values with highlighted lines showing the first timestep of each month. (b) The evolution of $\theta_b(\phi, q{=}2)$. (c) $\theta(\lambda, \phi, q{=}2)$ for the first forecast at 24-hours lead time. (d) θ' from Eq. 8: anomaly of (c) relative to the background state with $\theta'{=}0$ highlighted by the bold line. The white regions in (c) and (d) show the mask on $\theta(\lambda, \phi, q{=}2){>}340$ K.

In the midlatitudes, the equivalent latitude of the 2-PVU surface decreases monotonically going to higher θ surfaces (Fig. 1a). In this region a poleward displacement of the 2-PVU surface can be unambiguously associated with a positive θ anomaly (negative for an equatorward displacement). The exception is at the 340–350 K range corresponding to the subtropical jet: at the subtropical jet, the background state 2-PVU surface can be folded so that $\theta_b(\phi, q{=}2)$ is multivalued. *Chagnon and Gray* [2015] noted that the dipole of diabatically-generated PV across the 2-PVU surface was not robust in subtropical regions which is consistent with the tropopause equatorward of the subtropical jet not being well defined as a constant PV surface [*Wilcox et al.*, 2012]; therefore, regions where the forecast $\theta(\lambda, \phi, q{=}2)$ is greater than 340 K are excluded from the diagnostics calculated here. The background state $\theta_b(\phi, q{=}2, t)$ is then calculated by finding the θ that satisfies $\phi_e(\theta, q{=}2) = \phi$ by linear interpolation. In the case of multiple θ values, the value of θ less than 340 K is taken. Figure 1b shows $\theta_b(\phi, q{=}2, t)$. Note that there is no time averaging but that $\theta_b(\phi, q{=}2, t)$ is inherently slowly varying.

Figure 1c shows $\theta(\lambda, \phi, q{=}2)$ from the first forecast at 24-hours lead time and Fig. 1d shows the anomaly relative to the background state. Ridges and troughs are defined by the sign of the anomaly in Fig. 1d (positive and negative respectively). The advantage of this diagnostic is that it has allowed identification of ridges and troughs on a limited area domain even if it is much smaller than the scale of Rossby wave activity. The white regions in Fig. 1c and d show the mask applied at θ>340 K to ignore subtropical air masses. We find that there are occasionally regions of negative or near zero PV in the stratosphere associated with gravity-wave breaking that cause the tropopause to be diagnosed too high; the mask on θ>340 K is also useful for excluding these points.

4 Results

In this section the results from the winter season forecasts are presented. Composites of PV and PV tracer diagnostics relative to the tropopause are presented in section 4.1. In section 4.2 the tropopause-relative composites are used to quantify the evolution of tropopause sharpness with lead time and the contributions of different processes to tropopause sharpness. In the following sections the results from the first two sections are explained in terms of different processes: advection by the model winds (section 4.3), dynamics-tracer inconsistency (section 4.4) and parametrized physical processes (section 4.5).

4.1 Tropopause-relative composites

The novel method that led to the discovery of the TIL by *Birner et al.* [2002] was compositing radiosonde profiles relative to the diagnosed thermal tropopause. The composites in this study are produced in a coordinate relative to the dynamical tropopause, defined as the 2-PVU surface,

$$\tilde{z} = z - z(q{=}2). \tag{9}$$

The approach is similar to *Cavallo and Hakim* [2009] who used a coordinate of pressure relative to the tropopause to composite PV tendencies in tropopause polar vortices. The composites are produced using the following method:

1. For each forecast, at each lead time
   (a) Calculate the height of the 2-PVU surface using linear interpolation from PV on model levels. For any columns with multiple heights for the 2-PVU surface (i.e. folded tropopause), the highest position is taken.
   (b) Linearly interpolate each variable to height levels relative to the dynamical tropopause ($\tilde{z}$). The levels are taken every 0.2 km up to ±2 km from the tropopause. Note that this resolution is sharper than the vertical model grid spacing which decreases from 400 m at 6 km to 600 m at 12.5 km.
   (c) Calculate the area-weighted mean of each variable on each tropopause-relative level over areas diagnosed as ridges and troughs separately.
2. Calculate the mean and standard error of each diagnostic over the set of forecasts.

The compositing method above is then repeated taking $\tilde{z}$ relative to the 2-PVU surface of the advection-only PV tracer (q_{adv}=2) rather than q=2 in Eq. 9. Repeating the composites relative to each surface (q=2 and q_{adv}=2) allows us to systematically quantify how much non-conservative processes act on either side of the tropopause (the composites are the same) or directly influence stratosphere-troposphere exchange by separating the two surfaces (the composites are different). This can be seen if we consider some non-conservative process producing negative PV tendencies initially above the tropopause. In this case, initially stratospheric air (q>2) can become tropospheric (q<2) such that the diagnosed position of the

q=2 surface has moved above the negative PV tendencies but the position of the q_{adv}=2 surface is unchanged. The opposite can occur for positive PV tendencies initially below the tropopause with the position of the q=2 surface moving below the positive PV tendencies. Over many of these situations we would diagnose positive PV tendencies systematically above the q=2 surface but below the q_{adv}=2 surface and negative PV tendencies systematically below the q=2 surface but above the q_{adv}=2 surface. Therefore, a composite over many cases would diagnose dipoles across the q=2 and q_{adv}=2 surfaces of opposite sign; however, since the q=2 surface has moved this does not necessarily imply any change in the diagnosed PV gradient across the q=2 surface, only that mass is being exchanged between the troposphere and stratosphere. This would not be the case for positive PV tendencies above the tropopause or negative PV tendencies below the tropopause because they would not directly move the q=2 surface, and therefore have a direct effect on the PV gradient. This is also true of PV tendencies occurring near the tropopause that are too weak to directly move the q=2 surface. In these cases, the composites relative to q=2 and q_{adv}=2 would be the same and imply a direct effect on the tropopause PV gradient.

Figure 2 shows the tropopause-relative composites over ridges (a,b,c) and troughs (d,e,f) at 24-hours lead time. Figure 2a and d show PV and q_{adv} as the difference between the 24-hour forecasts and the verifying analyses for each forecast. The profile of PV in Fig. 2 is thus the systematic forecast error. There is a systematic decrease in PV above the 2-PVU surface relative to analyses, but comparatively little change in the troposphere (the error is zero at the tropopause because q=2 by definition). The systematic errors in PV can be contrasted with q_{adv} which reduces above the tropopause and increases below the tropopause relative to the analyses (Fig. 2a and d). The difference between PV and q_{adv} is the "net effect of non-conservative processes" ($q - q_{adv}$) which was shown to enhance the tropopause PV gradient by *Chagnon et al.* [2013]. The tendency of $q - q_{adv}$ is systematically positive in the stratosphere and negative in the troposphere (Fig. 2b and e) consistent with the case studies from *Chagnon et al.* [2013] and *Chagnon and Gray* [2015]. The effects of non-conservative processes are also of similar magnitude to the systematic forecast errors.

The PV tracers partition $q - q_{adv}$ (Eq. 7) into parametrized physical processes ($\sum q_{phys}$), dynamics-tracer inconsistency (ε_I) and a residual (ε_r). Figure 2b and e show that the residual is small with approximately zero systematic effect allowing us to focus on the remaining terms. The combined effect of parametrized physical processes ($\sum q_{phys}$) is to produce a dipole in PV tendencies with positive PV tendencies in the stratosphere and negative PV tendencies in the troposphere and approximately zero net change at the 2-PVU surface, consistent with the findings of *Chagnon et al.* [2013] and *Chagnon and Gray* [2015] from individual case studies. The dipole is similar when composited relative to q_{adv}=2, albeit weaker, showing that the parametrized physical processes are acting to directly enhance the tropopause PV gradient rather than change the height of the tropopause. The partitioning of $\sum q_{phys}$ into individual physical processes (Fig 2c and f) is discussed in section 4.5.

The dynamics-tracer inconsistency (ε_I) shows net negative tendencies at tropopause level in ridges and troughs (Fig. 2b and e) although there are positive PV tendencies around 1 km above the tropopause which are more pronounced in ridges than in troughs. The negative peak is slightly below q=2, but above q_{adv}=2, which indicates that, unlike the parametrized physical processes, the main effect of ε_I is to directly separate the two surfaces (q_{adv}=2 and q=2). This does not explain why ε_I is most negative at the tropopause which is discussed in section 4.4.

4.2 Tropopause PV contrast

To quantify the effects of different physical processes on the reduction in isentropic PV gradient seen in [*Gray et al.*, 2014] we can calculate the vertical tropopause PV contrast of the variables in Fig. 2 over a fixed distance. The vertical gradient of PV will be much larger than the isentropic PV gradient due to the typical slope of the tropopause; therefore we quan-

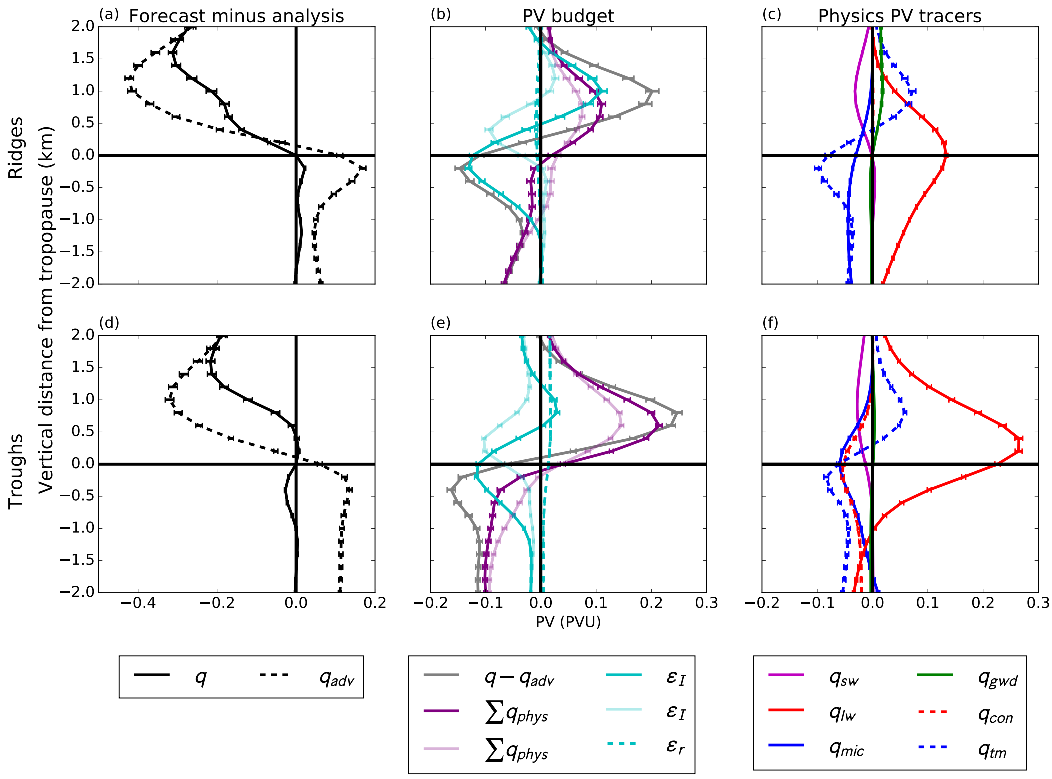

Figure 2. PV as a function of vertical distance from the 2-PVU surface in ridges (a, b and c) and troughs (d, e, and f). Lines show the mean and error bars show the standard error on the mean for the 92 forecasts at 24-hours lead time. (a) and (d) show the forecast minus analysis values for PV (q) and the advection-only PV tracer (q_{adv}). (b) and (e) show the difference ($q - q_{adv}$) and the contributing processes: parametrized physical processes ($\sum q_{phys}$), dynamics-tracer inconsistency (ε_I) and a residual (ε_r). The faint lines show composites relative to the advection-only PV tracer (q_{adv}=2) for ε_I and $\sum q_{phys}$. (e) and (f) show the contributions to $\sum q_{phys}$ from the individual physics tracers: short-wave radiation (q_{sw}), long-wave radiation (q_{lw}), microphysics (q_{mic}), gravity-wave drag (q_{gwd}), convection (q_{con}) and turbulent mixing (q_{tm}).

tify the PV contrast over a fixed vertical distance which is proportional to the isentropic contrast over a larger fixed horizontal distance if the tropopause slope is assumed to be constant. From the tropopause relative means, the tropopause PV contrast for each variable is calculated as the difference between the average of points 1 km above and below the tropopause. As with the previous composites, the mean and standard error are then calculated over the 92 forecasts.

Figure 3 shows the tropopause PV contrast as a function of lead time for each of the variables in Fig. 2. There is a reduction in PV contrast with lead time (Fig. 3a and d) consistent with the reduction in isentropic PV gradient found by *Gray et al.* [2014]. The reduction in PV contrast is stronger in ridges than in troughs.

The contrast in q_{adv} decreases more rapidly than for PV because it is not being maintained by diabatic processes: the parametrized physical processes produce a net increase in the tropopause PV contrast with lead time in ridges and troughs (Fig. 3b and e). The contribution of individual physical processes (Fig 3c and f) is discussed in section 4.5. The diagnosed contribution of ε_I to the tropopause PV contrast is less clear, showing an increased

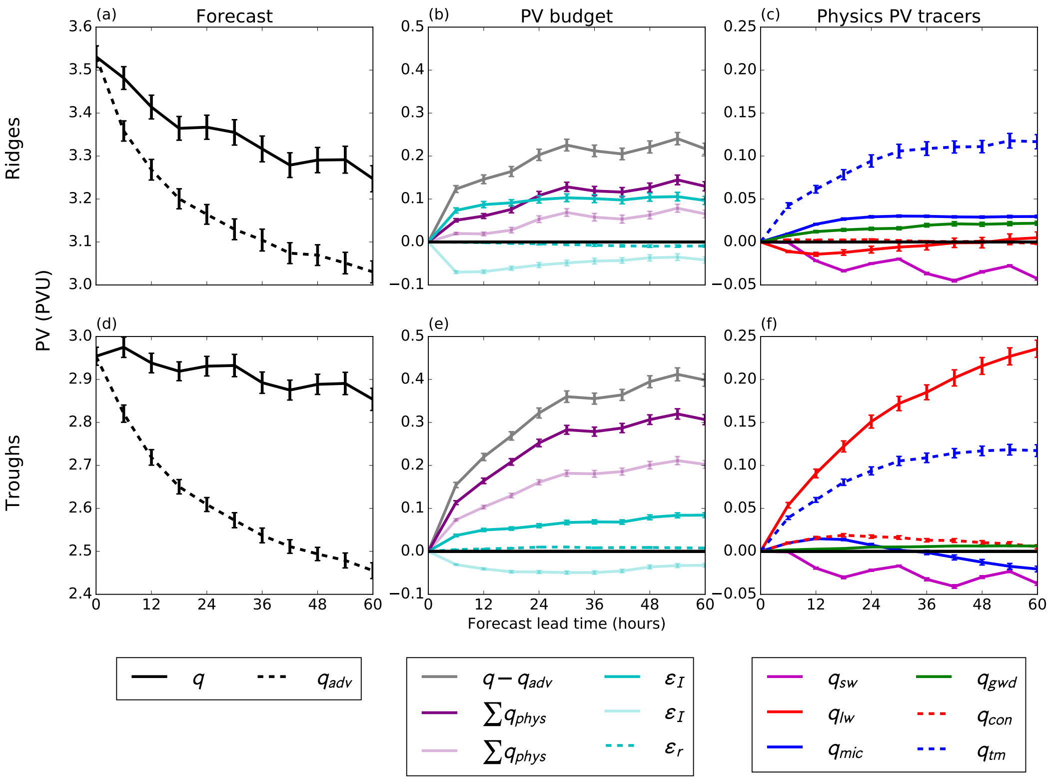

Figure 3. The same variables as in Fig. 2, but showing the tropopause PV contrast as a function of lead time calculated as the difference between points up to 1 km above and 1 km below 2-PVU. PV and the advection-only PV tracer are shown as absolute values rather than forecast minus analysis.

contrast relative to q=2 and a reduced contrast relative to q_{adv}=2. This is because, as stated in the previous section, ε_I is acting to directly separate the two surfaces.

4.3 Tracer advection

The evolution of q_{adv} is a result of advection by the resolved winds of the model using the semi-Lagrangian scheme of the MetUM. Conservative tracer advection results in a continuous cascade of features to smaller scales. Horizontal and vertical length scales in tracers decrease exponentially at the same rate [*Haynes and Anglade*, 1997] giving an exponential increase in tracer gradients. The difference here is that implicit numerical diffusion takes over as length scales approach the grid-scale and we are calculating the PV contrast over a fixed length scale. Diffusive processes act most rapidly at small scales and slowly at large scales. The contrast in q_{adv} decreases as features cascade to smaller scales where diffusion reduces the extrema.

The decrease of the contrast in q_{adv} is the opposite to that expected from vortex stripping (see section 2). The reason for this different behaviour is that there is a dynamical equilibrium between sharpening by intermittent stripping events and a continuous smoothing of the tropopause. A model with consistent initial conditions would be initialised in the dynamic equilibrium state of the model climate and the net effects of processes sharpening and smoothing the tropopause would cancel out over many forecasts, giving a constant PV gradient as a function of lead time. In the idealised simulations of *Ambaum* [1997] and *Haynes et al.* [2001] the diabatic processes contribute to a smoothing of the tropopause and so an advection-only PV tracer initialised in the dynamic equilibrium state would show a net sharpening of the tropopause on short timescales. In our simulations diabatic processes directly sharpen the tropopause so the net effect of the advection scheme must be to smooth the tropopause. Also, we start from an analysis in which gradients are sharper than can be maintained by the free-running model.

The net result of the tracer advection is that the contrast of q_{adv} as a function of lead time (T) exponentially decays from an initial contrast $\Delta q_{adv}(0)$, to a reduced contrast, $\Delta q_{adv}(\infty)$,

$$\Delta q_{adv}(T) = \Delta q_{adv}(\infty) + [\Delta q_{adv}(0) - \Delta q_{adv}(\infty)]e^{-\frac{T}{\tau}}, \tag{10}$$

where τ is the decay timescale. Although the term $\Delta q_{adv}(\infty)$ is obtained by fitting Eq. 10 to the forecast data, it cannot be a long-time limit for a passive tracer because a tracer will eventually become well mixed as diffusive effects dominate.

The parameters in Eq. 10 have been calculated by fitting Eq. 10 to the evolution of $\Delta q_{adv}(T)$ using scipy.optimize.curve_fit [*Jones et al.*]. Figure 4 shows an example of this fit for $\Delta q_{adv}(T)$ in ridges. The solid black line is the same as the dashed line in Fig. 3a and the grey line shows Δq_{adv} calculated from composites relative to q_{adv}=2 rather than relative to q=2. The evolution of Δq_{adv} relative to q_{adv}=2 is shown because the evolution can only be a result of the tracer advection scheme, even in the presence of non-conservative processes. The circles in Fig 4a show the fit of Eq. 10. Note that the first data point has been excluded from the fit leaving $\Delta q_{adv}(0)$ as a derived parameter. This was done because the first six-hours deviates slightly from an exponential decay. This can be seen from the fitted points: in the first six-hours Δq_{adv} decreases more rapidly relative to q=2 and slightly less rapidly relative to q_{adv}=2 compared to what would be predicted from the following exponential decay. The estimate of $\Delta q_{adv}(0)$ is not very sensitive to ignoring the first data point; however, the derived timescale is sensitive to overfitting to the first data point giving an overestimation of the timescale relative to q=2 and an underestimation of the timescale relative to q_{adv}=2.

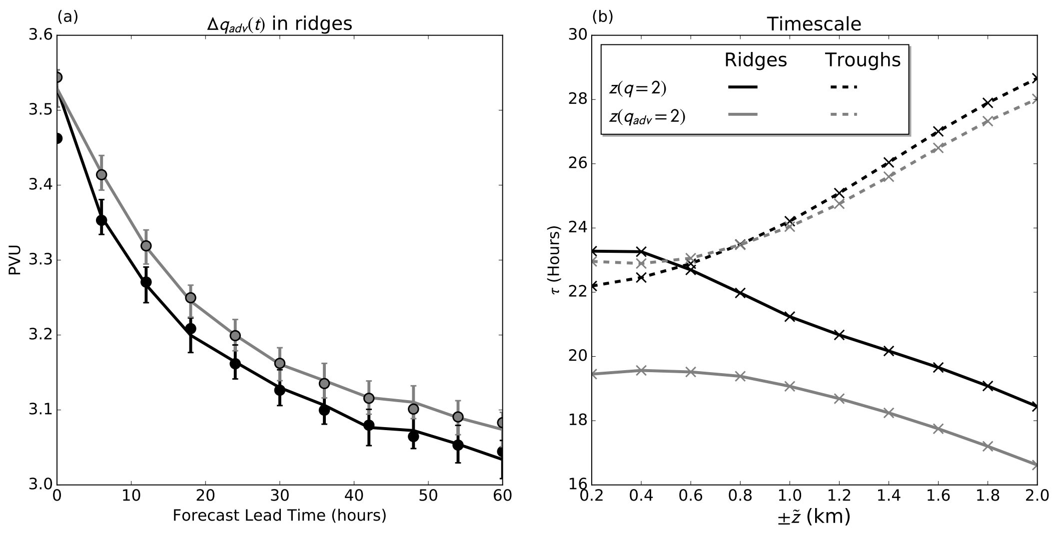

Figure 4. Fitting of Eq. 10 to the decay of Δq_{adv} in the forecasts. (a) Δq_{adv} in ridges (circles show fit). (b) The derived timescale from Eq. 10 for varying vertical scales. The key in (b) applies to both plots.

The fit of Eq. 10 is repeated for multiple vertical length scales by calculating Δq_{adv} only from points up to $\pm\tilde{z}$: Fig. 4b shows the derived timescale. The timescale appears to be constant at small vertical scales because we approach the vertical level spacing of the model which reduces from 400 m at 6 km to 600 m at 12.5 km altitude. Approaching the scale of the vertical resolution, the timescale increases in ridges and decreases at in troughs. It is unclear why the timescale in ridges and troughs should have the opposite behaviour as a function of vertical scale; however, the timescale does approach a similar value of 20−24 hours in ridges and troughs.

4.4 Dynamics-tracer inconsistency

The dynamics-tracer inconsistency quantifies the difference between non-conservation of PV resulting from the dynamical core and non-conservation associated with the tracer advection scheme. *Saffin et al.* [2016] showed that local tendencies of dynamics-tracer inconsistency were dominated by non-conservation of PV by the dynamical core; however, this result does not necessarily generalize to the integrated tendencies over many forecasts so it is important to diagnose the underlying processes.

Figure 5 shows the tropopause-relative mean of the dynamics-tracer inconsistency as a function of lead time. The top panels show composites relative to q=2 and the bottom panels show composites relative to q_{adv}=2. There is a dipole of positive and negative tendencies centered slightly above q=2 suggesting a raising and sharpening of the tropopause. However, the peak in negative tendencies is shifted upwards when composited relative to q_{adv}=2, as well as net positive tendencies appearing below q_{adv}=2 in ridges, as a result of the two surfaces (q_{adv}=2 and q=2) being separated by dynamics-tracer inconsistency rather than directly affecting the tropopause PV gradient (see section 4.1). The positive tendencies have saturated at short lead times which may explain the discrepancy in behaviour of Δq_{adv} over the first 6 hours in Fig. 4. This rapid saturation can also be seen for the diagnosed effect of dynamics-tracer inconsistency on the PV contrast in ridges (Fig. 3b).

At longer lead times the dynamics-tracer inconsistency becomes increasingly negative which is more pronounced in troughs. A possible explanation for the net negative tendencies of the dynamics-tracer inconsistency is that it results from dissipation as part of the vortex stripping process: as filaments of PV are drawn away from the tropopause the dynamical-core dissipates the PV filament faster than tracer advection giving negative tendencies in the filament. Negative PV tendencies are consistent with a downwards diabatic transport of mass by dilution of PV substance [*Haynes and McIntyre*, 1987]. This is consistent with the isentropic map of ε_I shown in *Saffin et al.* [2016] (their Fig. 2c) where net negative tendencies are seen in the troughs where q=2 is displaced from q_{adv}=2.

4.5 Parametrized physical processes

The combined effect of parametrized physical processes is to produce a dipole in PV tendencies with positive PV tendencies in the stratosphere and negative PV tendencies in the troposphere and zero net change at the 2-PVU surface (Fig. 2b and e), but this dipole is much weaker in ridges than troughs. These processes are now considered separately.

The largest contribution to the PV tendencies comes from the long-wave radiation which produces net positive PV tendencies at the tropopause and is about twice as strong in troughs as in ridges (Fig. 2c and f). Since the long-wave radiation is dependent on the humidity contrast and the absolute vorticity [*Forster and Wirth*, 2000], the stronger magnitude in troughs would be expected. Figure 6 shows variables from the analyses as a function of distance from the 2-PVU tropopause in ridges and troughs. Both the contrast in specific humidity (Fig. 6a) and the magnitude of the vertical component of the absolute vorticity (Fig. 6b) are approximately twice as strong in troughs as in ridges.

The contrast of the long-wave radiation PV tracer across the tropopause is also much stronger in troughs than ridges (Fig. 3c and f) which is due to the net PV tendencies being more symmetric across the tropopause in ridges than in troughs (Fig. 2c and f). The asymmetry of the net PV tendencies in troughs is likely due to the increased amount of clouds in troughs compared to ridges (Fig. 6c and d). As described in section 2, cloud-top cooling results in a sharp spike in diabatic cooling and, as a result, a dipole of PV tendencies. When

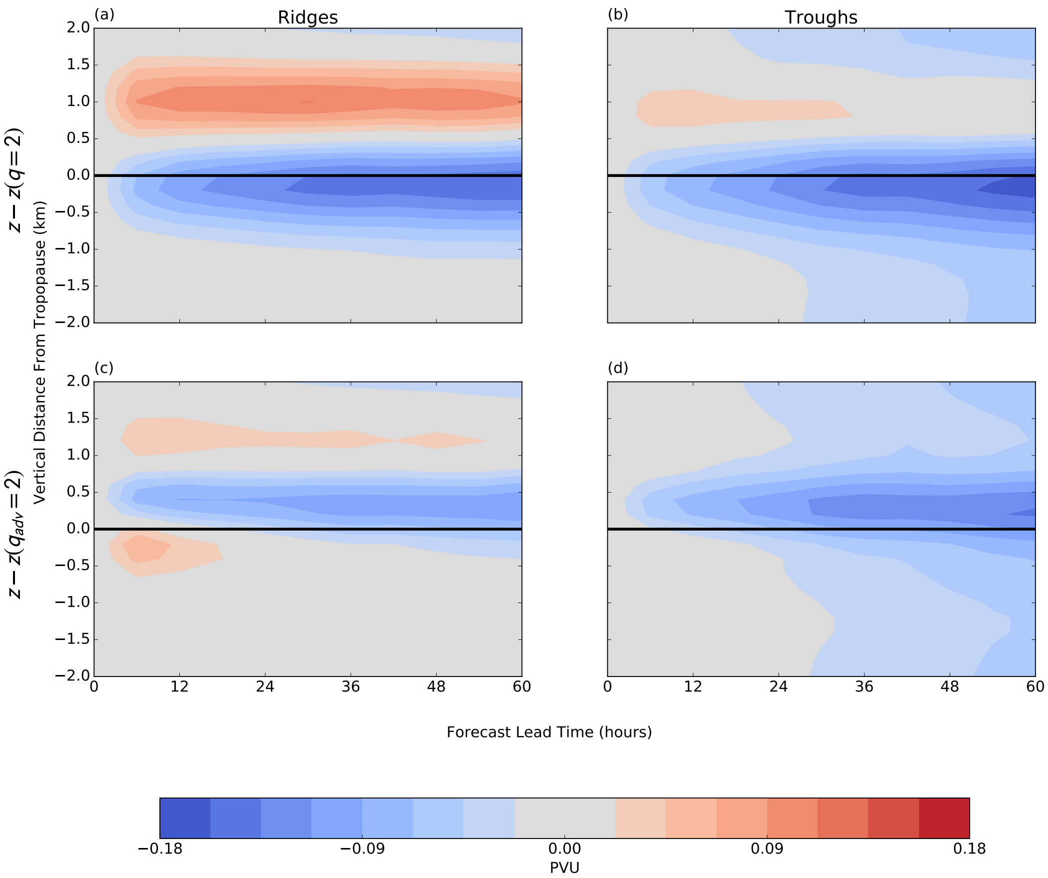

Figure 5. Dynamics-tracer inconsistency as a function of forecast lead time in ridges (a and c) and troughs (b and d). Values shown are the mean from the 3-months of forecasts when composited relative to the 2-PVU surface of PV (a and b) and the 2-PVU surface of the advection-only PV tracer (c and d)

composited over many clouds with varying distance from the tropopause this will show an enhanced gradient. *Cavallo and Hakim* [2009] showed that cloud-top cooling was a key process for intensifying tropopause polar vortices. The composites of PV tendencies relative to tropopause polar vortices from *Cavallo and Hakim* [2009] (their Fig. 9) show similar tendencies to those seen for our composite over troughs (Fig. 2) with net positive tendencies across the tropopause and negative tendencies further below the tropopause.

Short-wave radiation produces negative PV tendencies above the tropopause (Fig. 2c and f) which act to reduce the PV gradient with a clear diurnal cycle visible (Fig. 3c and f) since we are using a limited area domain. In both ridges and troughs, short-wave radiation reduces the PV gradient during the daytime by producing negative tendencies in PV above the tropopause. Negative PV tendencies indicate a negative heating gradient in the lower stratosphere which is most likely due to the variation in water vapour. Radiative heating due to ozone might be expected to have a large effect as positive PV tendencies below a heating maxima. Strongly positive values of the short-wave radiation PV tracer are seen at higher altitudes, but too far from the tropopause to affect the composites in Figs. 2 and 3.

The microphysics PV tracer shows a net negative PV accumulation below the tropopause in both ridges and troughs (Fig. 3c and f), consistent with the negative PV tendencies from ascent above the maxima in latent heating. The association of the turbulent-mixing PV tracer to vertical transport is less clear. *Chagnon et al.* [2013] associated negative values of the

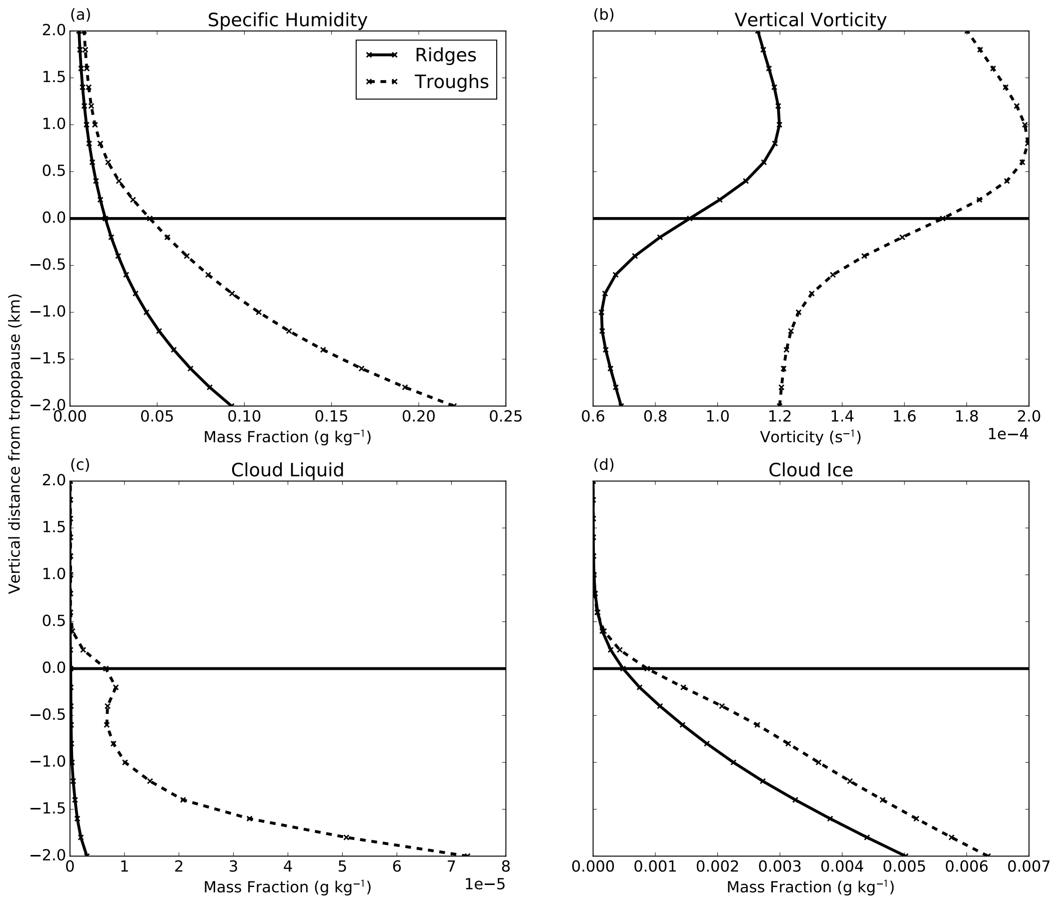

Figure 6. The mean of variables as a function of distance from the 2-PVU tropopause in ridges and troughs from the analyses used in this study. (a) Specific humidity, (b) vertical component of absolute vorticity, (c) mass fraction of cloud liquid and (d) mass fraction of cloud ice.

turbulent-mixing PV tracer with transport of tracer from the boundary-layer by a WCB. Ventilation of the boundary layer is dominated by WCBs [*Sinclair et al.*, 2008] so we might expect to see a signature of WCB transport to the tropopause in the turbulent-mixing PV tracer; however, we see an effect at short lead times, so it is important to distinguish between the effects of parametrized mixing at the tropopause-levels and long-range transport from the boundary-layer.

The diagnosed impact of processes related to WCBs depends on the length of the forecast. Figure. 7a and b show the tropopause relative mean of the microphysics PV tracer as a function of lead time. The microphysics PV tracer shows net negative values in both ridges and troughs at short lead times because we are sampling air in the region of negative PV tendencies above the latent heating maxima. At longer lead times ridges and troughs show quite different behaviour: in ridges the values of the microphysics PV tracer are consistently negative whereas in troughs the negative values of the microphysics PV tracer are gradually replaced with positive values. This is because the outflow of WCBs, where the net change in PV will be negative or zero, is typically associated with ridges whereas in troughs, where the tropopause is lower, we will be compositing over air masses that are still ascending or have been affected by latent heating that is not associated with WCBs.

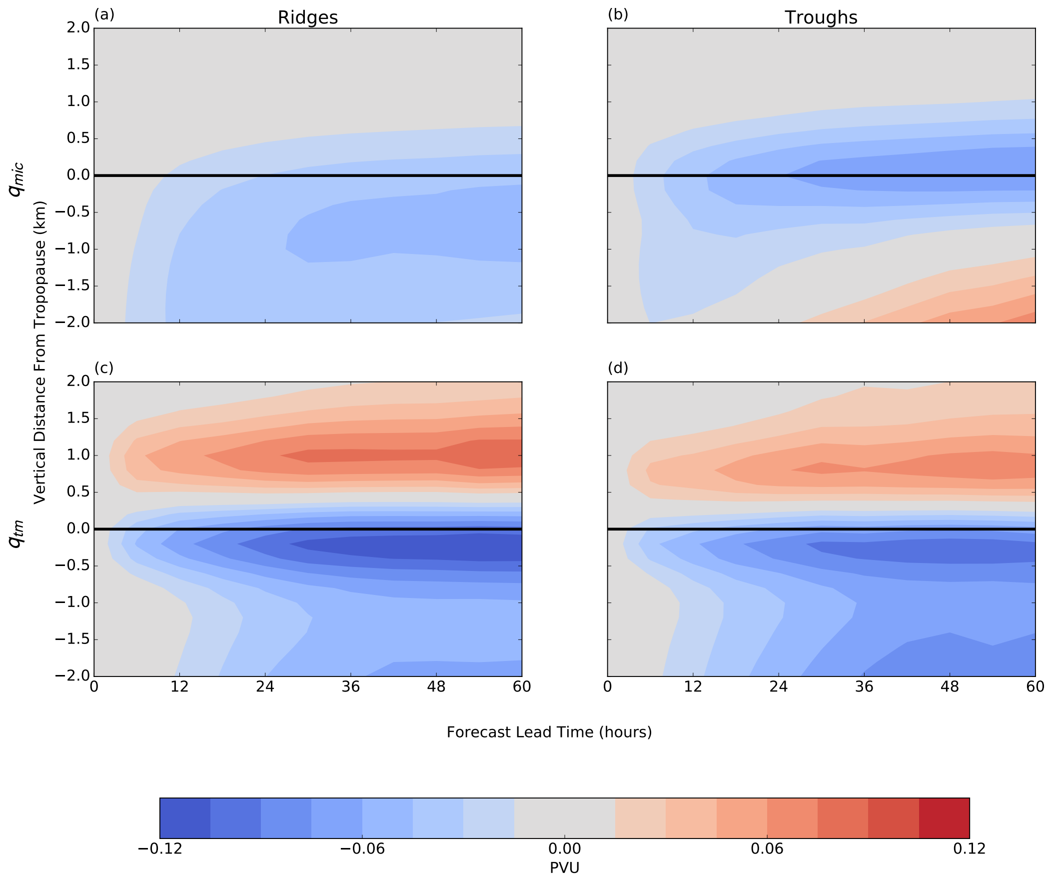

Figure 7. Microphysics (a and b) and turbulent mixing (c and d) PV tracers as a function of forecast lead time in ridges (a and c) and troughs (b and d). Values shown are the mean from the three-months of forecasts relative to the 2-PVU surface.

The turbulent mixing parametrization is having a systematic effect on the tropopause within the first six hours in both ridges and troughs which is unlikely to be from WCBs (Fig. 7c and d). The strongest negative tendencies are just below tropopause level and near zero further below indicating that they are not being advected from lower levels. There is a small hint of negative tendencies increasing from lower levels with lead time, but the dominant process is turbulent mixing at tropopause levels.

Convection has almost zero effect in ridges but shows net negative tendencies in troughs. This makes sense since the tropopause is lower in troughs and so we expect stronger convective transport in troughs. The reverse is true for gravity-wave drag with roughly net zero effect in troughs (Fig 2f) and a net positive in ridges (Fig 2c); however, this net positive is small and could be an artifact of large negative tendencies from gravity-wave drag causing the tropopause to be diagnosed too high which is then masked out in the composites (section 3.3).

5 Conclusions

Gray et al. [2014] showed that the tropopause PV gradient reduces with forecast lead time in NWP models; however, the source of model error remained unclear. In this study the systematic effects of individual model processes in maintaining the sharpness of the extratropical tropopause have been quantified by integrating PV tracers over a set of 92 forecasts with the MetUM. Since PV is conserved for adiabatic and frictionless flow [*Ertel*, 1942], PV tracers can accumulate tendencies of PV from individual model processes following the resolved flow which can be advantageous when compared to calculating Eulerian initial tendencies [*Klinker and Sardeshmukh*, 1992; *Rodwell and Palmer*, 2007] by avoiding large cancellations due to advection as will be the case near the tropopause. This study demonstrates that PV tracers can be a useful alternative to the initial tendencies method for quantifying the systematic behaviour of individual model processes.

Composites of PV tracers have been produced relative to the 2-PVU tropopause separately for ridges and troughs, diagnosed as anomalies of θ on the 2-PVU tropopause. Rossby waves are associated with meridional displacements of PV contours on isentropic surfaces which can be associated with an anomaly of θ on the 2-PVU tropopause.

The key results from this study are

1. The vertical PV contrast across the tropopause reduces relative to analyses with forecast lead time consistent with a smoothing of the isentropic PV gradient [*Gray et al.*, 2014].
2. On the timescales of the forecasts, the advection scheme of the model gives an exponential decay of the tropopause PV contrast to a finite value with a timescale of 20-24 hours.
3. A key component of the PV budget is the dynamics-tracer inconsistency which quantifies the difference between the evolution of PV in the dynamical core and the evolution of PV through tracer advection [*Saffin et al.*, 2016].
   (a) The locations of the maxima in dynamics-tracer inconsistency are different when composited relative to the 2-PVU surface of the advection-only PV tracer rather than PV indicating that dynamics-tracer inconsistency is having a direct effect on mass transport across the tropopause causing the q=2 and q_{adv}=2 surfaces to separate.
   (b) The dynamics-tracer inconsistency shows net negative tendencies near the tropopause level indicating a net transfer of mass from the stratosphere to the troposphere. This is consistent with numerical mixing removing small-scale stratospheric filaments and could be related to vortex stripping [*Ambaum*, 1997].
4. Parametrized physical processes act to sharpen the tropopause by producing a dipole in PV tendencies across the tropopause with near zero net tendency at the tropopause, consistent with the results of *Chagnon et al.* [2013] and *Chagnon and Gray* [2015] from individual case studies.
   (a) Radiative cooling produces net positive tendencies across and above the tropopause due to the gradient of water vapour across the tropopause. The stronger water vapour gradient and absolute vorticity in troughs compared to ridges results in a stronger net positive PV tendency. The positive PV tendencies due to radiative cooling have a stronger gradient across the tropopause in troughs than in ridges which can be explained by the increased frequency of clouds acting to sharpen the vertical cooling gradient [*Cau et al.*, 2005].
   (b) The microphysics PV tracer accumulates negative PV below the tropopause at short lead times associated with latent heating in WCBs.
   (c) The turbulent-mixing PV tracer accumulates negative PV at the tropopause and positive PV above the tropopause. At short lead times the majority of the turbulent mixing PV tracer seen at tropopause levels accumulates locally rather than by the long-range transport from the boundary layer seen at longer timescales.

The open question now is what changes should be made to NWP models to improve the representation of the tropopause PV gradient? The work in this paper provides a frame-work for testing such changes. We have used a limited area domain to give a resolution comparable to current global models at a lower computational cost; however, the inflow of air from the lateral boundaries results in an uncertainty in the behaviour of the tropopause as the PV tracers can not trace air prior to inflow. The recommendation for repeating this analysis to investigate model changes would be to use a global model but fewer forecasts.

An obvious first step would be to investigate changing model resolution. *Gray et al.* [2014] showed a reduction of PV gradient in day 10 of the ECMWF forecasts associated with a reduction in the horizontal resolution of the model. In this study we have shown a less dramatic reduction in tropopause sharpness than *Gray et al.* [2014] which is most likely because we have a higher resolution limited area domain ($0.11°$ here compared to a range between $0.28 - 1.4°$ for the forecasts analysed by *Gray et al.* [2014]). It would be useful to know if the model is accurately representing poorly resolved mixing processes. Varying the resolution of the forecasts should help identify errors arising from non-conservation of PV by the dynamical core and parametrized turbulence, both of which have been to shown to be important processes at tropopause levels.

We have shown that the magnitude of the systematic forecast error is comparable to the tendencies due to the parametrized physical processes so it is plausible that realistic modifications to the model parametrization schemes could significantly reduce the error rather than more difficult measures such as increasing resolution or redesigning the dynamical core. It would be expected that modifying the microphysics and/or convection schemes to enhance the latent heating driven ascent in WCBs would have a large impact on the tropopause. It is useful to consider how changes to WCBs would affect the PV tracers. The simplest effect would be to directly enhance the negative PV tendencies below the tropopause at short lead times. Increased transport of moisture and cloud formation at tropopause levels would also modify the response of radiative tendencies. We did not find any significant errors in forecast of water vapour or clouds near the tropopause when compared with analyses (not shown) which suggests the WCB transport in the forecasts is adequate; however, this raises the question of how much we trust the analyses. ECMWF analyses have been shown to have a moist bias in the lower stratosphere [*Dyroff et al.*, 2015] and the ERA-Interim reanalyses have been shown to have insufficient cloud and a low cloud-top bias in WCBs [*Hawcroft et al.*, 2016]. It is possible that an initial bias in the analyses is maintained through the forecasts leading to an underestimation of the effects of WCBs and long-wave radiation on the tropopause.

It is useful to associate systematic differences between forecasts and analyses with observations which is difficult for PV. The strength of the TIL [*Birner et al.*, 2002] is a useful measure of tropopause sharpness that can be obtained from observations. It is notable that studies on the processes affecting the TIL find similar results to studies on the processes affecting the tropopause PV gradient (e.g. *Chagnon et al.* [2013] and *Kunkel et al.* [2016]). Since static stability is proportional to the vertical gradient in θ, we would expect a region of enhanced static stability above the tropopause to be associated with a positive PV anomaly and a stronger PV gradient. *Pilch Kedzierski et al.* [2016] showed that, without data assimilation, the TIL region in the ECMWF NWP model tends towards a weaker static stability which is probably associated with the decline in PV gradient shown by *Gray et al.* [2014]; however, further work is needed to associate these two features. The recent North Atlantic Waveguide and Downstream Impact Experiment (NAWDEX) field campaign also provides an opportunity to compare observations and analyses of tropopause structure.

Acronyms

NWP Numerical weather prediction
MetUM Met Office Unified Model
PV Potential vorticity
PVU Potential vorticity units
TIL Tropopause inversion layer

WCB Warm conveyor-belt
NAE North Atlantic European (model configuration)

Notation

$\boldsymbol{\theta}$ Potential temperature
$\boldsymbol{q}$ PV
$\boldsymbol{q_{adv}}$ Advection-only PV tracer
$\boldsymbol{q_{lw}}$ Long-Wave Radiation PV tracer
$\boldsymbol{q_{sw}}$ Short-Wave Radiation PV tracer
$\boldsymbol{q_{mic}}$ Microphysics PV tracer
$\boldsymbol{q_{gwd}}$ Gravity-Wave Drag PV tracer
$\boldsymbol{q_{con}}$ Convection PV tracer
$\boldsymbol{q_{tm}}$ Turbulent-mixing PV tracer
$\boldsymbol{\varepsilon_I}$ Dynamics-tracer inconsistency

Acknowledgments

We thank the three anonymous reviewers for their detailed comments which have greatly improved this paper. The lead author is funded by a Natural Environment Research Council doctoral training grant (reference NE/L501608/1) and is grateful for CASE funding from the Met Office. K. D. Williams was supported by the Joint UK BEIS/Defra Met Office Hadley Centre Climate Programme (GA01101). We acknowledge use of the MONSooN system, a collaborative facility supplied under the Joint Weather and Climate Research Programme which is a strategic partnership between the Met Office and the Natural Environment Research Council. The edits required for the PV tracers are stored as a branch edit in PUMA's TRAC system. The branch used in this paper is /dev/LSaffin/vn7.3_PV_Tracers revision number 18699. The source code for the MetUM is available to use. To apply for a license for the MetUM go to http://www.metoffice.gov.uk/research/collaboration/um-collaboration. For more information on the exact model versions and branches applied please contact the authors. Data from the simulations is archived at the Met Office and available for research use through the Centre for Environmental Data Analysis JASMIN platform (http://www.jasmin.ac.uk/); please contact the authors for details.

References

Ambaum, M. H. P. (1997), Isentropic formation of the tropopause, *J. Atmos. Sci.*, *54*, 555–568, doi:10.1175/1520-0469(1997)054<0555:IFOTT>2.0.CO;2.

Arakawa, A., and V. R. Lamb (1977), Computational design of the basic dynamical processes of the UCLA general circulation model, *Methods Comput. Phys.*, *17*, 173–265.

Birner, T., A. Dörnbrack, and U. Schumann (2002), How sharp is the tropopause at midlatitudes?, *Geophys. Res. Lett.*, *29*(14), 44–45, doi:10.1029/2002GL015142.

Cau, P., J. Methven, and B. Hoskins (2005), Representation of dry tropical layers and their origins in ERA-40 data, *J. Geophys. Res. Atmos.*, *110*(D6), doi:10.1029/2004JD004928.

Cavallo, S. M., and G. J. Hakim (2009), Potential Vorticity Diagnosis of a Tropopause Polar Cyclone, *Mon. Weather Rev.*, *137*(4), 1358–1371, doi:10.1175/2008MWR2670.1.

Chagnon, J. M., and S. L. Gray (2015), A Diabatically Generated Potential Vorticity Structure near the Extratropical Tropopause in Three Simulated Extratropical Cyclones, *Mon. Weather Rev.*, *143*(6), 2337–2347, doi:10.1175/MWR-D-14-00092.1.

Chagnon, J. M., S. L. Gray, and J. Methven (2013), Diabatic processes modifying potential vorticity in a north atlantic cyclone, *Q. J. R. Meteorol. Soc.*, *139*(674), 1270–1282, doi:10.1002/qj.2037.

Charney, J. G., and N. A. Phillips (1953), Numerical Integration of the Quasi-Geostrophic Equations for Barotropic and Simple Baroclinic Flows, *J. Meteorol.*, *10*(2), 71–99, doi: 10.1175/1520-0469(1953)010<0071:NIOTQG>2.0.CO;2.

Davies, T. (2014), Lateral boundary conditions for limited area models, *Q. J. R. Meteorol. Soc.*, *140*, 185–196, doi:10.1002/qj.2127.

Davies, T., M. J. P. Cullen, A. J. Malcolm, M. H. Mawson, A. Staniforth, A. A. White, and N. Wood (2005), A new dynamical core for the Met Office's global and regional modelling of the atmosphere, *Q. J. R. Meteorol. Soc.*, *131*(608), 1759–1782, doi:10.1256/qj.04.101.

Davis, C. A., M. T. Stoelinga, and Y.-H. Kuo (1993), The Integrated Effect of Condensation in Numerical Simulations of Extratropical Cyclogenesis, *Mon. Weather Rev.*, *121*(8), 2309–2330, doi:10.1175/1520-0493(1993)121<2309:TIEOCI>2.0.CO;2.

Dee, D. P., S. M. Uppala, A. J. Simmons, P. Berrisford, P. Poli, S. Kobayashi, U. Andrae, M. A. Balmaseda, G. Balsamo, P. Bauer, P. Bechtold, A. C. M. Beljaars, L. van de Berg, J. Bidlot, N. Bormann, C. Delsol, R. Dragani, M. Fuentes, A. J. Geer, L. Haimberger, S. B. Healy, H. Hersbach, E. V. Hólm, L. Isaksen, P. Kållberg, M. Köhler, M. Matricardi, A. P. McNally, B. M. Monge-Sanz, J.-J. Morcrette, B.-K. Park, C. Peubey, P. de Rosnay, C. Tavolato, J.-N. Thépaut, and F. Vitart (2011), The ERA-Interim reanalysis: configuration and performance of the data assimilation system, *Q. J. R. Meteorol. Soc.*, *137*(656), 553–597, doi:10.1002/qj.828.

Dyroff, C., A. Zahn, E. Christner, R. Forbes, A. M. Tompkins, and P. F. J. van Velthoven (2015), Comparison of ECMWF analysis and forecast humidity data with CARIBIC upper troposphere and lower stratosphere observations, *Q. J. R. Meteorol. Soc.*, *141*(688), 833–844, doi:10.1002/qj.2400.

Edwards, J. M., and A. Slingo (1995), Studies with a flexible new radiation code. I: Choosing a configuration for a large-scale model, *Q. J. R. Meteorol. Soc.*, *122*(531), 689–719, doi: 10.1002/qj.49712253107.

Ertel, H. (1942), Ein neuer hydrodynamischer Wirbelsatz, *Meteorol. Zeitschrift*, *59*, 277 – 281.

Ferreira, A. P., J. M. Castanheira, and L. Gimeno (2016), Water vapour stratification and dynamical warming behind the sharpness of the Earth's midlatitude tropopause, *Q. J. R. Meteorol. Soc.*, *142*(695), 957–970, doi:10.1002/qj.2697.

Forster, C., and V. Wirth (2000), Radiative decay of idealized stratospheric filaments in the troposphere, *J. Geophys. Res. Atmos.*, *105*(D8), 10,169–10,184, doi: 10.1029/2000JD900052.

Grams, C. M., H. Wernli, M. Böttcher, J. Čampa, U. Corsmeier, S. C. Jones, J. H. Keller, C. J. Lenz, and L. Wiegand (2011), The key role of diabatic processes in modifying the upper-tropospheric wave guide: A North Atlantic case-study, *Q. J. R. Meteorol. Soc.*, *137*(661), 2174–2193, doi:10.1002/qj.891.

Gray, S. L. (2006), Mechanisms of midlatitude cross-tropopause transport using a potential vorticity budget approach, *J. Geophys. Res. Atmos.*, *111*(17), 1–14, doi: 10.1029/2005JD006259.

Gray, S. L., C. M. Dunning, J. Methven, G. Masato, and J. M. Chagnon (2014), Systematic model forecast error in Rossby wave structure, *Geophys. Res. Lett.*, *41*(8), 2979–2987, doi: 10.1002/2014GL059282.

Grazzini, F., and F. Vitart (2015), Atmospheric predictability and Rossby wave packets, *Q. J. R. Meteorol. Soc.*, *141*(692), 2793–2802, doi:10.1002/qj.2564.

Gregory, D., and P. R. Rowntree (1990), A Mass Flux Convection Scheme with Representation of Cloud Ensemble Characteristics and Stability-Dependent Closure, *Mon. Weather Rev.*, *118*(7), 1483–1506, doi:10.1175/1520-0493(1990)118<1483:AMFCSW>2.0.CO;2.

Harrold, T. W. (1973), Mechanisms influencing the distribution of precipitation within baroclinic disturbances, *Q. J. R. Meteorol. Soc.*, *99*(420), 232–251, doi: 10.1002/qj.49709942003.

Harvey, B. J., J. Methven, and M. H. P. Ambaum (2016), Rossby wave propagation on potential vorticity fronts with finite width, *J. Fluid Mech.*, *794*, 775–797.

Hawcroft, M., H. Dacre, R. Forbes, K. Hodges, L. Shaffrey, and T. Stein (2016), Using satellite and reanalysis data to evaluate the representation of latent heating in extratropical cyclones in a climate model, *Clim. Dyn.*, pp. 1–24, doi:10.1007/s00382-016-3204-6.

Haynes, P., and J. Anglade (1997), The Vertical-Scale Cascade in Atmospheric Tracers due to Large-Scale Differential Advection, *J. Atmos. Sci.*, *54*(9), 1121–1136, doi:10.1175/1520-0469(1997)054<1121:TVSCIA>2.0.CO;2.

Haynes, P., J. Scinocca, and M. Greenslade (2001), Formation and maintenance of the extratropical tropopause by baroclinic eddies, *Geophys. Res. Lett.*, *28*(22), 4179–4182, doi:10.1029/2001GL013485.

Haynes, P. H., and M. E. McIntyre (1987), On the Evolution of Vorticity and Potential Vorticity in the Presence of Diabatic Heating and Frictional or Other Forces, *J. Atmos. Sci.*, *44*(5), 828–841, doi:10.1175/1520-0469(1987)044<0828:OTEOVA>2.0.CO;2.

Hoskins, B., and P. Berrisford (1988), A Potential Vorticity Perspective of the Storm of 15-16 October 1987, *Weather*, *43*(3), 122–129, doi:10.1002/j.1477-8696.1988.tb03890.x.

Hoskins, B. J., and T. Ambrizzi (1993), Rossby Wave Propagation on a Realistic Longitudinally Varying Flow, *J. Atmos. Sci.*, *50*(12), 1661–1671, doi:10.1175/1520-0469(1993)050<1661:RWPOAR>2.0.CO;2.

Hoskins, B. J., M. E. McIntyre, and A. W. Robertson (1985), On the use and significance of isentropic potential vorticity maps, *Q. J. R. Meteorol. Soc.*, *111*(470), 877–946, doi:10.1002/qj.49711147002.

Jones, E., T. Oliphant, P. Peterson, and Others (), SciPy: Open source scientific tools for Python.

Joos, H., and H. Wernli (2012), Influence of microphysical processes on the potential vorticity development in a warm conveyor belt: a case-study with the limited-area model COSMO, *Q. J. R. Meteorol. Soc.*, *138*(663), 407–418, doi:10.1002/qj.934.

Klinker, E., and P. D. Sardeshmukh (1992), The Diagnosis of Mechanical Dissipation in the Atmosphere from Large-Scale Balance Requirements, *J. Atmos. Sci.*, *49*(7), 608–627, doi:10.1175/1520-0469(1992)049<0608:TDOMDI>2.0.CO;2.

Kunkel, D., P. Hoor, and V. Wirth (2016), The tropopause inversion layer in baroclinic life-cycle experiments: the role of diabatic processes, *Atmos. Chem. Phys.*, *16*(2), 541–560, doi:10.5194/acp-16-541-2016.

Legras, B., and D. Dritschel (1993), Vortex stripping and the generation of high vorticity gradients in two-dimensional flows, *Appl. Sci. Res.*, *51*(1), 445–455, doi:10.1007/BF01082574.

Lock, A. P., A. R. Brown, M. R. Bush, G. M. Martin, and R. N. B. Smith (2000), A New Boundary Layer Mixing Scheme. Part I: Scheme Description and Single-Column Model Tests, *Mon. Weather Rev.*, *128*(9), 3187–3199, doi:10.1175/1520-0493(2000)128<3187:ANBLMS>2.0.CO;2.

Madonna, E., H. Wernli, H. Joos, and O. Martius (2014b), Warm Conveyor Belts in the ERA-Interim Dataset (1979-2010). Part I: Climatology and Potential Vorticity Evolution, *J. Clim.*, *27*(1), 3–26, doi:10.1175/JCLI-D-12-00720.1.

Martius, O., C. Schwierz, and H. C. Davies (2010), Tropopause-Level Waveguides, *J. Atmos. Sci.*, *67*(3), 866–879, doi:10.1175/2009JAS2995.1.

Methven, J. (2015), Potential vorticity in warm conveyor belt outflow, *Q. J. R. Meteorol. Soc.*, *141*(689), 1065–1071, doi:10.1002/qj.2393.

Methven, J., and P. Berrisford (2015), The slowly evolving background state of the atmosphere, *Q. J. R. Meteorol. Soc.*, *141*(691), 2237–2258, doi:10.1002/qj.2518.

Palmer, T. N., F. J. Doblas-Reyes, A. Weisheimer, and M. J. Rodwell (2008), Toward Seamless Prediction: Calibration of Climate Change Projections Using Seasonal Forecasts, *Bull. Am. Meteorol. Soc.*, *89*(4), 459–470, doi:10.1175/BAMS-89-4-459.

Pilch Kedzierski, R., L. Neef, and K. Matthes (2016), Tropopause sharpening by data assimilation, *Geophys. Res. Lett.*, *43*(15), 8298–8305, doi:10.1002/2016GL069936.

Randel, W. J., F. Wu, and P. Forster (2007), The Extratropical Tropopause Inversion Layer: Global Observations with GPS Data, and a Radiative Forcing Mechanism, *J. Atmos. Sci.*, *64*(12), 4489–4496, doi:10.1175/2007JAS2412.1.

Riemer, M., and S. C. Jones (2010), The downstream impact of tropical cyclones on a developing baroclinic wave in idealized scenarios of extratropical transition, *Q. J. R. Meteorol. Soc.*, *136*(648), 617–637, doi:10.1002/qj.605.

Rodwell, M. J., and T. N. Palmer (2007), Using numerical weather prediction to assess climate models, *Q. J. R. Meteorol. Soc.*, *133*(622), 129–146, doi:10.1002/qj.23.

Saffin, L., J. Methven, and S. L. Gray (2016), The non-conservation of potential vorticity by a dynamical core compared with the effects of parametrized physical processes, *Q. J. R. Meteorol. Soc.*, *142*(696), 1265–1275, doi:10.1002/qj.2729.

Scaife, A. A., N. Butchart, C. D. Warner, and R. Swinbank (2002), Impact of a Spectral Gravity Wave Parameterization on the Stratosphere in the Met Office Unified Model, *J. Atmos. Sci.*, *59*(9), 1473–1489, doi:10.1175/1520-0469(2002)059<1473:IOASGW>2.0.CO;2.

Schemm, S., H. Wernli, and L. Papritz (2013), Warm Conveyor Belts in Idealized Moist Baroclinic Wave Simulations, *J. Atmos. Sci.*, *70*(2), 627–652, doi:10.1175/JAS-D-12-0147.1.

Sinclair, V. A., S. L. Gray, and S. E. Belcher (2008), Boundary-layer ventilation by baroclinic life cycles, *Q. J. R. Meteorol. Soc.*, *134*(635), 1409–1424, doi:DOI 10.1002/qj.293.

Son, S.-W., and L. M. Polvani (2007), Dynamical formation of an extra-tropical tropopause inversion layer in a relatively simple general circulation model, *Geophys. Res. Lett.*, *34*(17), doi:10.1029/2007GL030564.

Thuburn, J. (2008), Some conservation issues for the dynamical cores of NWP and climate models, *J. Comput. Phys.*, *227*(7), 3715–3730, doi: http://dx.doi.org/10.1016/j.jcp.2006.08.016.

Wang, S. M., and M. A. Geller (2016), Baroclinic mixing of potential vorticity as the principal sharpening mechanism for the extratropical Tropopause Inversion Layer, *Earth Sp. Sci.*, *3*(9), 362–369, doi:10.1002/2015EA000150.

Webster, S., A. R. Brown, D. R. Cameron, and C. P.Jones (2003), Improvements to the representation of orography in the Met Office Unified Model, *Q. J. R. Meteorol. Soc.*, *129*(591), 1989–2010, doi:10.1256/qj.02.133.

Wernli, H., and H. C. Davies (1997), A lagrangian-based analysis of extratropical cyclones. I: The method and some applications, *Q. J. R. Meteorol. Soc.*, *123*(538), 467–489, doi: 10.1002/qj.49712353811.

Whitehead, J. P., C. Jablonowski, J. Kent, and R. B. Rood (2015), Potential vorticity: Measuring consistency between GCM dynamical cores and tracer advection schemes, *Q. J. R. Meteorol. Soc.*, *141*(688), 739–751, doi:10.1002/qj.2389.

Wilcox, L. J., B. J. Hoskins, and K. P. Shine (2012), A global blended tropopause based on ERA data. Part I: Climatology, *Q. J. R. Meteorol. Soc.*, *138*(664), 561–575, doi: 10.1002/qj.951.

Wilson, D. R., and S. P. Ballard (1999), A microphysically based precipitation scheme for the UK meteorological office unified model, *Q. J. R. Meteorol. Soc.*, *125*(557), 1607–1636, doi:10.1002/qj.49712555707.

Wirth, V. (2001), Cyclone-Anticyclone Asymmetry Concerning the Height of the Thermal and the Dynamical Tropopause, *J. Atmos. Sci.*, *58*(1), 26–37, doi:10.1175/1520-0469(2001)058<0026:CAACTH>2.0.CO;2.